DOCUMENT RESUME

ED 274 962 CS 008 576

AUTHOR Malliet, Gineva M.
TITLE The Effect of Colour-Word Interference on Children's Memory for Words.
PUB DATE Aug 86
NOTE 48p.; Ph.M. Thesis, University of Cambridge.
PUB TYPE Reports - Research/Technical (143) -- Dissertations/Theses - Master Theses (042)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Processes; *Color; Comparative Analysis; Elementary Education; Foreign Countries; Instructional Materials; Pretests Posttests; *Printing; Reading Processes; *Reading Research; *Recall (Psychology); Retention (Psychology); Short Term Memory
IDENTIFIERS *Stroop Color Word Test

ABSTRACT

The Stroop color-word test involves a conflict situation in which subjects are asked to say aloud the ink color used to print a color word on a card. Interference occurs when the ink color is in conflict with the color word, such as 'red' printed in green ink. On the other hand, little interference occurs when asked to name the color words themselves, regardless of ink color. A study examined this interference effect on children's immediate retention of words printed in color. Subjects, 70 boys and girls, 8 and 9 years old, were pretested for memory ability on a criterion list of 20 neutral words printed in black and white, and assigned to one of three stimulus conditions. In the first, subjects were presented with color words printed in related colors; in the second, subjects were presented with the same color words, but printed in unrelated colors; and in the third, subjects were presented with the same color word printed in black and white. Results of the immediate recall test indicated that the scores for the related color group were superior to those of the unrelated color group, but only for memory as opposed to correct spelling. Mean scores for the unrelated group were also the lowest of the three conditions for both memory and spelling. No other significant differences were detected between the three groups. Results did not support assumptions that the motivational benefits of color in educational materials automatically lead to increased learning. An eight-page bibliography and five data tables are appended. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED274962

The Effect of Colour-Word Interference on Children's Memory for Words

by Gineva M. Malliet

A Dissertation Submitted in Partial Fulfilment of the Requirements of the Master of Philosophy Degree

University of Cambridge

August, 1986

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gineva M. Malliet

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

CS 008576

TABLE OF CONTENTS

ABSTRACT.....2

THE EFFECT OF COLOUR-WORD INTERFERENCE ON CHILDREN'S MEMORY FOR WORDS

INTRODUCTION.....3

Words in Colour.....4

The Stroop Colour-Word Effect.....7

Colour Research.....9

METHOD.....14

Subjects and Design.....14

Materials.....16

Procedure.....18

Scoring.....18

RESULTS.....19

DISCUSSION.....21

BIBLIOGRAPHY.....28

TABLE 1.....37

TABLE 2.....38

TABLE 3.....39

TABLE 4.....40

TABLE 5.....41

FIGURE CAPTIONS.....42

FIGURE 1

FIGURE 2

The vowels of the English language always appear to me, when I think of them, as possessing certain colours.... Consonants, when thought of by themselves, are of a purplish black; but when I think of a whole word, the colour of the consonants tends towards the colour of the vowels. For example, in the word "Tuesday," when I think of each letter separately, the consonants are purplish black, *u* is a light dove colour, *e* is a pale emerald green, and *a* is yellow; but when I think of the whole word together, the first part is a light gray-green, and the latter part yellow. Each word is a distinct whole. I have always associated the same colours with the letters, and no effort will change the colour of one letter, transferring it to another. Thus the word "red" assumes a light green tint, while the word "yellow" is light green at the beginning and red at the end. Occasionally, when uncertain how a word should be spelt, I have considered what colour it ought to be, and have decided in that way. I believe this has often been a great help to me in spelling, both in English and foreign languages.... I cannot recall to mind anything that should have first caused me to associate colours with letters, nor can my mother remember any alphabet or reading book coloured in the way I have described which I might have used as a child.

Letter to Sir Francis Galton, published in his *Inquiries Into Human Faculty*, 1883.

The Effect of Colour-Word Interference on Children's Memory for Words

Gineva M. Malliet

University of Cambridge

Abstract

A colour-word interference task was used to test the effect of that interference on children's immediate retention of words. Interference was found in this variation of the Stroop test in that the related colour group performed significantly better than the unrelated colour group on the memory task. Results are discussed in terms of both previous investigations into levels of processing in the Stroop phenomenon, and studies of the motivational or organizational benefits of colour in educational materials. The conclusion is reached that the level of processing for colour in children is perhaps deeper than has been previously recognized. Only qualified support for colour's organizational utility was found, and results do not support assumptions that the motivational benefits of colour in educational materials automatically lead to increased learning.

The Effect of Colour-Word Interference

on Children's Memory for Words

Colour is often used in the printing of children's books, and in reading schemes and other educational materials designed for both children and adults. While the use of colour can increase the cost of such materials tenfold, such use is generally justified by colour's assumed aesthetic, attentional, or motivational value (e.g., Moyle & Moyle, 1974; Southgate & Roberts, 1970; Watts & Nisbet, 1974). Claims are also made for the usefulness of colour as a mnemonic device which makes the material more meaningful and aids in the retention of material to be learned (cf. Watts & Nisbet, 1974). While some evidence exists for the motivational value of colour, the slight research into the educational value of colour in general often produces inconclusive results or results which mitigate colour's motivational utility (cf. Otto & Askov, 1968). Such research and its findings will be discussed later in this paper.

In a slightly different vein, colour as it relates to the Stroop colour-word effect has been extensively researched since 1935, when the phenomenon was first described by Stroop. These studies generally have focused on the attentional and automatic processing aspects of cognition revealed by the phenomenon. Little experimentation has been done, however, which directly applies the Stroop phenomenon to memory and learning tasks.

The experiment described in this paper attempts to explore this underexamined aspect of the Stroop colour-word effect by testing the memory aspects of the phenomenon to discover what those aspects might reveal with regard to practical questions of

educational materials currently in use.

In this introduction, one of the reading schemes which relies on colour as a motivational/learning aid will first be discussed, followed by an examination of experiments using the Stroop colour-word effect which impinge upon the results of the current experiment.

Words in Colour

Printed media used in the teaching of reading has traditionally been adapted in one of two ways: Either the spelling of the words is changed to resemble speech sounds, as in, for example, i.t.a., or some kind of signalling system is used. In the 18th and 19th centuries, diacritical marking was the popular signalling system; since then, the use of colour has come into favour.

Several colour reading schemes have been devised since N. Dale introduced colour as a word recognition aid in the 1890s (e.g., *Reading by Rainbow* and *Colour Story Reading*), but in general they fall into two categories: either colour is used to signal a particular phonic rule, or the phonic sounds are classified by colour (Moyle & Moyle, 1974).

Gattegno (1962, 1969) was the first person to devise a colour coding system designed to encompass all the phonic possibilities of the English language. His system, called *Words in Colour*, uses 47 different colours, each colour representing one sound, regardless of the spelling of that sound.

As with all the colour systems mentioned, Gattegno's system is held not only to regularize the complexities of English, but

to enhance attention and motivation. While it could be argued that an array of 47 different colours is more likely to confuse than motivate, there is evidence that learners prefer coloured materials, and that this preference may influence attention or motivation (e.g., Casey, 1972/1973; Daniel & Tacker, 1974; DeLucia, 1975; Jones, 1965). In discussing Chute s (1979, 1980) conclusions that colour has an adverse effect because it diverts attention from important learning cues, Lamberski and Dwyer (1983) argue that this adverse effect is highly conditional on the requirements of the specific task, and that colour may be important in providing an associative structure (pp. 18-19).

However, while the value of initial attraction to the learning materials is not inconsequential, it is generally accepted that the content of a book or lesson, rather than colour, is the primary motivational factor (D'Arcy, 1973). It should also be mentioned that sex differences can be a confounding factor, since great differences in colour preferences exist between girls and boys (Child, Hansen, & Hornbeck, 1968).

The motivation argument is perhaps best summed up by the results of Smith and Watkins' (1972) experiment, in which random colouring of test cards of danger objects in the kitchen was shown to be more attractive to children than black and white, but that comprehension scores were higher for black and white test cards.

The Gattegno system can also be criticized for reasons not having to do with questions of motivation. The following is a summary of points against the use of colour in the Gattegno system:

1) Learning sets are created through the use of the system which can hinder transfer of training (cf. Turner, 1984).

2) Because of the number of colours used, many of the colours are very close to each other in shade (e.g., tan, buff, and light brown), which can cause perceptual difficulties.

3) While the 47 colours used need not be memorized, some time is involved in learning the coding system. Gattegno (1962), in writing of the code's use in teaching languages such as Spanish, claims mastery of the code in fewer than ten lessons for 'ordinary adults,' with similar results for children (p. 23). However, ten lessons can be assumed to involve the expenditure of a large amount of classroom time; this needs to be justified by unequivocal results.

4) The system is, of course, useless for colour-blind students. [Jones (1967a) reports that about eight per cent of males have some form of defective colour vision, but that total monochromism occurs in only one person in a million. Dvorine (1962) puts the estimate at ten per cent.]

5) Children entering school, especially working class children, are not always familiar with colours, failing such tasks as colour matching or colour identification (Hunt & Kirk, 1974).

6) The use of such a code is inconsistent with the Piagetian theory that a child must have reached the stage of concrete operations before being able to decode (Moyle & Moyle, 1974).

7) A child using the Gattegno system, or any other regularized system, would experience difficulty upon moving to a new school that does not use the system.

8) Legibility can be a problem, since some colours are far more 'readable' than others (Harrison, 1980; Tinker, 1963).

Confusion also exists in the research as to whether or not colour preferences are in conflict with legibility; however, the familiarity of black and white print is always the determining factor in the legibility of print (Watts & Nisbet, 1974, p. 76).

The value of Gattegno's system, then, is often felt to be due more to the comprehensiveness of his systematization of language than to the use of colour *per se* (cf. McHugh, 1968). For example, Lockmiller and Di Nello (1970) found no significant differences between results with *Words in Color* and a basal reader method. The question seems to be not so much whether colour should be used at all in reading schemes, in light of its presumed motivational benefits, but rather how much colour should be used, and at what point this use becomes detrimental. Investigation of the Stroop effect can do much to enhance our knowledge of these points.

The Stroop Colour-Word Effect

The Stroop colour-word test involves a conflict situation in which subjects are asked to say aloud the ink colour used to print a colour word on a card. Interference occurs when the ink colour is in conflict with the colour word, e.g., 'red' printed in green ink. On the other hand, when subjects are asked to name patches of colour, their reaction times are much faster, and when asked to read colour words printed in black, their reactions are

faster still. In contrast, little interference occurs when subjects are asked to name the colour words themselves when those words are printed in incongruent colours: interference is so small, in fact, compared to black and white, that the difference could only be the result of the slightly reduced legibility of the words printed in colour (Dyer, 1972, p. 4).

This is an extraordinarily reliable phenomenon: although many experiments using variations of the Stroop test have been conducted, the results are highly uniform. The phenomenon is little understood, although it is clear that even when subjects are asked to ignore the word printed on the card, automatic processing of that word still occurs. Much discussion centers on the depth and rate of processing involved; the reason the phenomenon has attracted so much attention, as Dyer (1972) points out, is that it provides 'an important means for increasing our understanding of the normal processes of reading, stimulus identification and stimulus naming that either disrupt or are disrupted in the Stroop test' (p. 1).

Investigating the phenomenon further, Klein (1964) showed that interference also occurs when subjects are asked to name the ink colour for colour-related words, e.g., 'grass' printed in blue ink. His results indicated that interference also increases with the meaningfulness of the words, i.e., nonsense syllables cause less interference than meaningful words, which cause less interference than colour-related words.

Apparently the interference demonstrated by the Stroop test occurs at a deep level of analysis, with both the attended and

the unattended stimuli being processed to a semantic level. It is the speed of processing for the stimuli which seems to differ, with word processing occurring at a more rapid rate than colour processing (Farisse, 1969).

The discussion of research into colour which follows is noticeably lacking in references to the Stroop phenomenon as applied to learning situations; i.e., while various attempts have been made to investigate the practicality of the use of colour in learning and memory situations, these attempts have rarely explored the use of realistic or congruent colour versus incongruent colour.

Colour Research

As mentioned above, one of the main questions raised by the use of colour in reading schemes is one of transfer. Once children learn, for example, that the long 'a' sound in 'bake' is always coloured red in a particular reading scheme, will they experience difficulty when encountering that letter printed in the customary black and white?

Goodman and Cundick (1976) found, in a test involving the use of 14 modified Hebrew letters, that children aged six to nine learned faster with coloured learning materials. However, the children experienced difficulty during the transfer task to black and white. Lyczak (1976) obtained similar results with graduate students in a test involving 16 Thai syllables. The conclusion reached in this study, which also involved the variables of letter-shape training versus colour code training, was that letter shapes are not automatically memorized, and that training in the forms of letters is more crucial for learning than

training in the use of a colour code.

Similarly, Doyle (1982) found that using colour-coded letters as an aid to remediating reversals in words containing the letters p, b, and d was not as effective as simply practising the same words printed entirely in black and white.

Turner (1984) found that a colour code of 13 colours, for vowels only, facilitated the learning of irregularly spelled or nonphonetic words, but that these nonphonetic words were also harder to transfer to black than regular words. However, since all the children in her study were first trained using black and white materials, that training undoubtedly influenced both the colour and black and white groups in her study.

In experiments directly addressing the question of colour's effect on memory, colour is generally not found to have long-term benefits, although Winters and Hoats (1984) found that words isolated by colour in a sentence-like format facilitated immediate (rather than long-term) recall of those words in 36 retarded and nonretarded persons of similar mental age. The groups tested here were so small, though, as to place the results in doubt.

Farley and Grant (1976), in an experiment involving a black and white versus a colour slide photograph presentation, also found that on the immediate retention test, the colour group performed better. However, when both groups were tested seven days later, it was found that colour only slightly aided recall.

Lamberski and Dwyer (1981) also used slides (in this case

consisting of line drawings and topical outlines identifying the components of a video recorder system) in black and white or colour to test for immediate and delayed retention. They found that in three out of five of the immediate criterion measures, the black and white cueing technique was superior to colour. When subjects were tested after two weeks, no significant differences between the colour and black and white groups were found. The conclusion was reached that their data provide support for 'the contention that the mere addition of color in visualized instruction does not automatically produce increases in student achievement' (pp. 127-130), and that the effectiveness of colour decreases as the level of cognitive task increases. In other words, colour codes seem to be most useful in low-level cognitive tasks, such as sorting, when externally paced materials are used (Lamberski & Dwyer, 1983).

The number of colour-coded items used in an experiment can also be a confounding factor. Kanner (1968) reports that as the number of colour-coded items in a visual display increases, the usefulness of colour as a cueing agent decreases.

In 1982, Berry and Dwyer devised an experiment involving 224 college students which parallels the aims of the present experiment. They wished to test whether realistic or unrealistic colour, used in a slide presentation of drawings of the human heart, would have an effect on delayed retention. (Unrealistic colour was created through a process of photographic reversal that changed the realistic colour to its opposite or complementary colour.)

It was found that the usefulness of realistic or

nonrealistic colour was a function of IQ: Students of lower IQ did as well with black and white visuals as with colour; students of medium IQ were helped by the realistic colour, and generally were not helped by the use of unrealistic colour; and students of high IQ used unrealistic colour as easily as the other students used realistic colour cues. The conclusion was reached that 1) extreme realism in instructional visuals does not aid students of lower ability and 2) unrealistic colour can increase learning only for students of higher intelligence. Overall, the use of black and white visuals was found to be best for reducing individual differences associated with ability. And, in line with the results of other experiments mentioned above, neither colour realism, nor learners' IQ, affected retention after six weeks (Berry & Dwyer, 1982, p. 1091).

Lamberski and Dwyer (1983) conducted a further experiment which did lend support to the hypothesis that colour-coded instructional books would improve scores on immediate and delayed retention tests. For example, in an illustration of the aorta, the color gold was used for

1) the verbal illustration label AORTA

2) the arrow leading from AORTA to the visual area representing the structure of the aorta

3) the visualized shaded area indicating the aorta

4) the word AORTA cited twice in the accompanying prose text (p. 12).

This use of colour did facilitate retention on both the immediate and six-week delayed testing, even though the delayed

retention scores for the colour group declined more rapidly than for the black and white group (27 per cent decline in mean achievement scores versus 24 per cent decline). They concluded that the primary value of the colour code, as used here, was in providing 'a structure for physical discrimination and associative meaning in self-paced instruction,' while also pointing out that this conclusion is 'applicable to older learners and should not necessarily be generalized to younger learners' (p. 19).

The experiment described in this paper attempts to determine, as in Berry and Dwyer's 1982 experiment using photographic colour reversal, if colour enhances retention and, if so, under what colour conditions: related or unrelated. It differs from most of the experiments surveyed above in that 1) words rather than visual displays or photographs are used and 2) the effects of the Stroop colour-word phenomenon regarding colour conflicts are integrated into the experiment.

The question to be decided, then, is whether colour *per se* enhances retention, and is of some value in instructional materials, or whether only that colour which is inherently and semantically associated with the material to be learned is of any value in those materials, as the Stroop phenonmenon would seem to indicate.

Results of previous experiments involving the Stroop phenomenon show clearly that processing of words and colour can be highly contingent upon the association between the two items. However, since little interference occurs when subjects are asked to name colour words (rather than ink colour) when those words are printed in unrelated colours, we might expect that,

while there might be differences in memory between related and unrelated colour conditions, those differences will not prove to be significant. Compared to black and white, however, we might expect that semantically associated colour will improve memory because of children's demonstrated preference for colour.

Specifically, the hypotheses to be tested are as follows:

H1) Colour which is logically associated with words to be learned does not significantly enhance retention of those words compared to colour unrelated to those words.

H2) Colour, whether related or unrelated to words to be learned, enhances performance compared to black and white.

Method

Subjects and Design

The 70 subjects were 35 boys and 35 girls ranging in chronological age from 8 years, 1 month to 9 years, 7 months at the time of the experiment. The mean age for boys was 9 years, 3 months; the mean age for girls was 9 years, 2 months. All were enrolled in either Form 3, 4, or 5 in a housing estate school in a lower- to middle-income neighborhood.

Subjects' chronological ages and reading ages were obtained from school records. The children had been tested for reading age six months prior to the experiment using the Salford Test A: scores on this test ranged from 6 years, 5 months to 13 years, 3 months. The mean reading age was 9 years, 1 month for boys and 9 years, 4 months for girls.

Each child was pretested for memory ability on a criterion list of twenty neutral words printed in black and white. (See

Figure 1.) Words for this test and all tests used in the experiment were taken from compiled lists of words commonly used by children in this age range (Edwards & Gibbon, 1973; McNally & Murray, 1962). No attempt was made to standardize either the length of words or the number of syllables contained in each word.

Each form contained some children of extremely low ability: their teachers were consulted beforehand to determine which students would find the memorization task either impossible or anxiety-provoking. On the teacher's advice, these students were either not tested (i.e., testing of the remaining subjects was carried out while these students were absent), or their scores were removed from those of the final population before statistical analyses were performed. Similarly, the scores of two colour blind students were not included for analysis.

On the basis of performance on the black and white pretest, subjects were assigned to three stimulus conditions, so that each condition contained students drawn from the entire range of memorization ability. Scores on the black and white pretest were also used in some of the later statistical analyses, e.g., correlation of reading age with memory. While no deliberate attempt was made to counterbalance the three conditions by age or reading scores, these scores tended nonetheless to be equitably represented in each condition. Where possible, an approximately equal number of boys and girls was placed in each condition.

In the first stimulus condition, subjects (N=23) were presented with colour words printed in related colours; in the second condition, subjects (N=23) were presented with the same

colour words printed in unrelated colours; in the third condition, subjects (N=24) were presented with the same colour words printed in black and white. Again, no attempt was made to standardize either the length of the words or the number of syllables contained in the words. This procedure was necessary because few words exist in the language which are consistently associated with only one colour. (See Figure 2).

Analysis showed, however, that any differences in degree of difficulty between the words on the pretest and words on the subsequent stimulus condition tests were not significant, i.e., subjects performed almost equally well on the pretest as on the black and white stimulus condition test. This would seem to indicate that although the two lists of words were not scientifically matched for degree of difficulty, the lists were nonetheless roughly interchangeable.

Subjects, during both the pretest and in each stimulus condition, were given three minutes to study and attempt to memorize their respective lists. After studying the lists for three minutes, subjects were given an immediate retention test in which they were asked to write down, again in three minutes, as many of the words as they could recall.

No attention was drawn to the fact that colour was being used in two of the conditions, or that the colour in those conditions was either related or unrelated to the colour word.

Materials

Four sets of materials were created, each set consisting of strips of white paper 11 and 3/4 inches by 2 and 3/4 inches.

Each strip of paper contained one vertical row of twenty hand printed words. The materials were reproduced by either a standard photocopying method, in the case of the pretest list, or by a colour photocopying method, in the case of the three stimulus condition lists, so that each child was given his or her own individual list to study during the three-minute study phase of the test. The paper used in the standard photocopying method was non-glossy; that used in the colour photocopying method was heavier, with a glossy finish.

Six colours were used in both the related and unrelated colour conditions: red (used 4 times in both the related and unrelated colour lists), blue (4x), green (4x), brown (4x), orange (2x), and purple (2x). The colours otherwise appeared in random order on the lists, with the restriction that a word printed in one colour never appeared adjacent to another word printed in the same colour. Words were hand printed in lower case, since lower case is much more legible than all-capital printing (Tinker, 1963). Commercially available felt tip pens were used.

For conditions in which colour was not used (pretest and black and white stimulus condition), words were hand printed in black ink, again using a felt tip pen of the same type used for the colour conditions.

For the immediate retention test (given after both the pretest and the stimulus condition tests), subjects were asked to write down as many words from the lists as they could remember, using their own pencils and standard white, lined notepaper.

Procedure

Each subject was tested in the classroom setting at his or her usual desk.

Subjects were informed before the pretest that they would be given a list of words to try to memorize to the best of their abilities. They were asked, in addition, to pay close attention to the spellings of the words. They were then informed that after being given three minutes to study the list of twenty words, they would be asked to write down, again in three minutes, as many of the words as they could recall. Again, it was mentioned that they should attempt to reproduce the spellings of the words as exactly as possible.

To reduce the anxiety inherent in a test situation, however, the children were told that the test was to be regarded as a game or a contest, rather than an examination.

The strips of paper containing the test words were then handed face down to the subjects, who were instructed not to turn the papers over until given the signal to start.

The three stimulus condition tests were given one week after the pretest; the same instructions were again given beforehand, and the identical procedure was followed. At no time during the experiment was colour mentioned, nor was the children's attention drawn to the colour used in two of the conditions.

Scoring

Scoring was divided into two parts for both the pretest and for the three stimulus conditions. Subjects were graded on 1) perfect recall of a word in the list, i.e., for perfect spelling of that word, and 2) memory, even if imperfect, of a word. For

both types of scoring, one point was given for a correct answer. As a result, scores for memory were naturally uniformly higher than scores for spelling in subsequent analyses. This division in scoring was necessary because such young subjects were used in the experiment, for while it was clear in many cases that the subject could recall having seen a word on the list, he or she lacked the ability to spell that word other than phonetically (e.g., 'kitten' spelled as 'kittn'). Thus, scoring was conducted in such a way that children who were already good readers and spellers were not automatically given the advantage when scores were computed, and a more accurate reflection of memory improvement could be obtained for all the children tested.

Results

Unrelated t tests were performed to determine differences in performance between each of the three stimulus conditions. The t tests showed that scores for the related colour group were superior to scores for the unrelated colour group at the .05 level of confidence, but only for the memory (as opposed to the spelling) scoring method. Mean scores for the unrelated group were also the lowest of the three conditions for both memory and spelling [mean = 8.5652 versus 10.4783 (related) and 9.2083 (black and white) for memory; mean = 6.6522 versus 7.4348 (related) and 7.6667 (black and white) for spelling.] Aside from this finding of significant differences between the related and unrelated colour groups for memory, no significant differences were detected between the related, unrelated, and black and white groups for either memory or spelling. (See Table 1.)

Unrelated t tests were performed to test for sex differences in performance for both memory and spelling scores on the pretest, related colour, unrelated colour, and black and white tests. On the test using related colour, no significant differences were found between males and females; however, the girls did significantly better on 1) the pretest for both memory and spelling, 2) the black and white stimulus condition test for memory, and 3) the unrelated colour test for spelling, even though boys in this last condition were older (mean = 9 years, 5 months for boys versus 9 years for girls). Similar findings of sex differences studies showing that girls perform better on reading tasks have been well documented in the literature (e.g., Goodacre, 1971). However, in this last condition, the F value obtained was large (F value = 5.49; 2-tail probability = 0.021), suggesting that there was significantly more variation in the girls' scores than in the boys'.

Overall, girls performed better on all but the related colour test for spelling (mean = 7.7273 for boys versus 7.1667 for girls), although these differences in means on the related colour test were not significant. All the above-mentioned significant t test results were significant at the .05 level of confidence. (See Tables 2 and 3.)

The Spearman rank correlation test method was used to determine how well chronological age and reading age would correlate with performance on all tests, including the pretest. Chronological age did not correlate significantly with performance in any of the conditions. (See Table 4.) However, significant correlations at the .001 level of confidence were

found for reading age. Apparently, as reading age goes up, so does the ability to memorize this type of material, regardless of the congruity or incongruity of the colours used. (See Table 5.)

This finding also ties in with the above finding that girls performed better than boys on most of the tests, since the girls' average reading ages were three months higher than the boys'. It would appear, then, that both sex and reading age are factors in the ability to suppress the conflict in the colour-word interference phenomenon.

Applying these findings to the learning-to-read situation, a prediction can be made that poor readers might have more difficulty in coping with words printed in a wide array of colours than with black and white print; further research using the 'Gattegno-type' print, i.e., each sound printed in a different colour, needs to be carried out to confirm this hypothesis. However, the findings seem to coincide with the Berry and Dwyer (1982) study into the usefulness of realistic or unrealistic colour as a function of IQ, where students of high IQ were not confused by the use of unrealistic colour. While IQ scores were not used in the present experiment, here the students of higher reading ability also seemed better able to gate the incongruent colours.

Discussion

Results for the t test for memory comparing the related with the unrelated colour groups impinge on both colour-word interference studies and applications of colour to reading. Even though students were given more than sufficient time to suppress

the Stroop conflict in the unrelated colour condition, that conflict still appeared in the memory scores. This effect occurred even though only four words in the list were 'true' colour name words (discounting violet and orange, which are ambiguous words); interference has been shown to decrease sharply for colour words (e.g., grass) printed in incongruent colours (Hochman, 1969). Thus, H1 was not upheld in the case of the memory scores.

While the t test findings regarding related versus unrelated colour for the spelling scores were not significant, mean performance for the related group was still higher (mean = 7.4348 for related versus mean = 6.6522 for unrelated).

These findings seem to refute Kroll, Kellicutt, Berrian, and Kreisler's (1974) study of 32 undergraduates concerning what dimensions of an image are stored, whether stored intentionally or unintentionally. This experiment used test letters in upper and lower case or red or green colour. They hypothesized that colour might remain an integral part of a stored picture, regardless of whether or not that information regarding colour is relevant at the time of retrieval, but their findings led them to suggest that as retention intervals become longer (up to 8 seconds), the memory image is colourless unless subjects know that they will later be asked to report that colour, in which case the memory stores a coloured picture for later retrieval. Evidence for low incidental memory for print colour was also found by Bradlyn and Rollins (1980).

The subjects in the present study, however, had an average of nine seconds to study each word, and were not asked to notice

or remember the colours of the words, and yet it is clear from the scores of the unrelated colour group that the colour of the image was being stored, thus producing a memory conflict. The findings here do not support the idea of a colourless memory image, for while it is likely that the print colour was not as deeply analyzed as the word itself, some substantial degree of colour processing was clearly taking place.

There are two possible reasons why the findings here conflict with those of Kroll et al: 1) the difference in the age groups studied, and 2) the fact that only two colours were used in the Kroll et al study. It is possible that if colour does have the organizational benefit claimed by Lamberski and Dwyer (1983), two colours are not sufficient to aid in physical discrimination in a series of test letters.

Klatzky discusses a similar study of incidental memory for letter case and colour conducted by Light and Berger (1976), in which the two image attributes were found to be retained independently. He argues that 'the stored representation cannot be like a picture of a coloured word, which would tie the colour of the item to its form by representing the two together. The independent storage of the two attributes...is consistent with a more abstract representation (Klatzky, 1980, pp. 205-206).

The results of the present experiment also conflict with those found by Park and James (1983) in their study of 144 first, third, and fifth graders regarding incidental retention of colour information. They found that colour memory for pictures of simple objects in red or green was poorly retained by all age

groups unless the subjects tried to learn the colour. (Again, the fact that only two colours were used here may be of significance to the organizational issue.) While they found evidence for automatic processing of colour (i.e., attempting to learn colour did not impede memory for the objects), they suggest that colour is more an effortful than automatic process in retention, since their colour-intentional groups performed better in colour memory scoring than their colour-incidental groups. While incidental memory for colour was not assessed in the present experiment, the findings here nonetheless provide more support for an automatic than an effortful model in the case of a colour-word interference task. Evidence of automatic processing is also provided by Hale and Piper (1973), who found that children of different ages had a relatively high incidental memory for the colour of shapes they were instructed to learn.

Bradlyn and Rollins (1980), in discussing the results of their experiment testing incidental memory for colour-word association in the Stroop colour-word test, suggest that subjects have some control over attention to print colour, so that when processing of print colour might cause interference, that aspect of the stimulus can be gated before interference can occur (p. 272). Their experiment was conducted with undergraduate students, however; the present experiment does not support such a conclusion for children. They go on to state that performance in a colour-word interference task may be a function of the ability of subjects to gate print colour; the findings of the present experiment uphold that idea, suggesting that such ability is probably developmental.

The results here may coincide with Hagen's (1967) conclusions that ability to attend selectively to central information is a function of age. In his study with first, third, fifth, and seventh graders (using picture pairs rather than colour, however), it was the older rather than younger children in the study who could focus their attention on the relevant central information and ignore the incidental information. Clearly, the present experiment needs to be repeated with children from a wider range of age groups to explore further the possible developmental differences involved.

However, in the present experiment, reading age was a far more important indicator than chronological age: future developmental studies of verbal learning might do well to use reading age, rather than chronological age, as the criterion.

Of interest to the motivational issue is the fact that in the spelling scores the scores for the black and white group were highest (mean = 7.6667) -- even slightly higher than for the related colour group -- which tends to mitigate previously discussed findings concerning colour's motivational value. It is possible that the superior legibility of black and white was involved in this finding, but it is clear that, in this instance, H2 was not upheld. Even for the memory scores, where the black and white group did not do as well as the related colour group, the difference between the related colour group (mean = 10.4783) and the black and white group (mean = 9.2083) was not significant enough (1-tail probability = 0.114) to give unqualified support to the idea that colour was of motivational benefit.

The fact that the spelling scores for the black and white group were higher than for either related or unrelated colour also impinges on assumptions that colour serves an organizational function. In the design of this experiment, colours were repeated more than once; for example, red was used four times (for the words red, tomato, strawberry, and cherry). Assuming that children in the related colour group noticed that some words were of the same colour, they could have used that fact, deliberately or unconsciously, as a mnemonic aid during recall, which would have improved their scores over the black and white group. The fact that this was not reflected in the spelling scores does not necessarily denigrate colour's organizational benefit, since for the memory scores, in which recall of the word, however imperfect, was all that was required, the colour does seem to have been of some benefit.

Overall, however, the differences found here for the related colour and black and white groups are not large enough to justify the time and expense of using colour in educational materials for either assumed motivational or organizational purposes. Had the children in the related colour group first been instructed in the organizational use of the colour, i.e., had it been pointed out to them that they could group the words by colour, the results might have been different. This remains an area requiring further experimentation. Also, a truer test of the motivational value of related colour would be obtained in an experimental design in which colours were not repeated and thus were not available for mnemonic purposes; i.e., an experiment in which a different colour was used for each word.

Because of constraints necessarily placed on a limited study of this sort, several other fields of exploration suggested by the findings here were left unexamined. The experiment should be repeated using adults, to eliminate problems posed by the memory versus spelling design used here. Incidental memory for unrelated colours should be tested to investigate depth of processing of colour. It would also be useful to discover how much the time element is a factor, particularly with different age groups. Finally, possible cross-cultural differences in colour memory should be investigated.

While it is clear that colour has an effect on memory in the colour-word interference task described, this effect is not sufficiently significant to warrant the inclusion of colour in an educational setting. Even though results for the related colour versus unrelated colour situation were significant, this finding is not practically generalizable to education except in highly specific contexts, and is more of value in indicating an area for further study applying the Stroop phenomenon to memory research.

The assumed motivational value of colour is not upheld by the study, nor is conclusive evidence found for the organizational utility of colour. Further systematic research is needed in the many areas already described to provide a comprehensive assessment of colour's utility for children of all levels of ability.

Bibliography

Acredolo, L., Fick, H., & Olsen, M. (1975). Environmental differentiation and familiarity as determinants of children's memory for spatial location. *Developmental Psychology, 11,* 495-501.

Babbitt, B. (1982). Effect of task demands on dual coding of pictorial stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 73-80.

Baddeley, A. (1983). *Your memory.* New York: Penguin.

Berry, L., & Dwyer, F. (1982). Interactive effects of colour realism and learners' IQ on effectiveness of visual instruction. *Perceptual and Motor Skills, 54,* 1087-1091.

Birren, F. (1950). *Color psychology and color therapy.* New York: McGraw-Hill.

Bradlyn, A., & Rollins, H. Jr. (1980). Incidental memory for the color-word association in the Stroop color-word test. *Bulletin of the Psychonomic Society, 16,* 269-272.

Casey, R. Jr. (1973). The cognitive effectiveness of color and monochrome educational materials with elementary school children (Doctoral dissertation, University of Southern California, 1972). *Dissertation Abstracts International, 33,* 1072A.

Child, I., Hansen, J., & Hornbeck, F. (1968). Age and sex differences in children's colour preferences. *Child Development, 39,* 237-47.

Chute, A. (1979). Analysis of the instructional functions of color and monochrome cueing in media presentations. *Educational Communications and Technology, 27,* 251-263.

Chute, A. (1980). Effect of color and monochrome versions of a film on incidental and task-relevant learning. *Educational Communications and Technology, 28,* 10-18.

Dalrymple-Alford, E. (1972). Associative facilitation and interference in the Stroop color-word task. *Perceptual Psychophysics, 11,* 274-276.

Daniel, P., & Tacker, R. (1974). Preferred modality of stimulus input and memory for CVC trigrams. *Journal of Educational Research, 67,* 255-256.

D'Arcy, P. (1973). *Reading for meaning 2: The reader's response.* London: Hutchinson Educational.

Dean, J. (1966). Words in colour. In Downing, J. (ed.) *The first international reading symposium, Oxford 1964.* London: Cassell.

DelCastillo, D., & Gumenik, W. (1972). Sequential memory for familiar and unfamiliar forms. *Journal of Experimental Psychology, 9,* 90-96.

DeLucia, F. (1975). Comparative effectiveness of monochrome and color presentations in the facilitation of affective learning by low-income black elementary school pupils (Doctoral dissertation, University of Southern California). *Dissertation Abstracts International, 35,* 5673A.

Diack, H. (1960). *Reading and the psychology of perception.* Nottingham, England: P. Skinner.

Doyle, W. (1982). The effectiveness of color-coded cues in remediating reversals. *Journal of Learning Disabilities, 15*, 227-230.

Dvorine, I. (1962). What is color blindness? *Education, 83*, 174-178.

Dwyer, F. (1972). *A guide for improving visualized instruction.* State College, PA: Learning Services.

Dwyer, F. (1978). *Strategies for improving visual learning.* State College, PA: Learning Services.

Dyer, F. (1971). Color-naming interference in monolinguals and bilinguals. *Journal of Verbal Learning and Verbal Behavior, 10*, 297-302.

Dyer, F. (1972). *The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes.* Fort Knox, KY: Army Medical Research Laboratory.

Edwards, R., & Gibbon, V. (1973). *Words your children use.* London: Burke Books.

Farisse, P. (1969). Why is naming longer than reading? *Acta Psychologica, 30*, 96-103.

Farley, F., & Grant, A. (1976). Arousal and cognition: Memory for color versus black and white multimedia presentation. *The Journal of Psychology, 94*, 147-150.

Galton, F. (1883). *Inquiries into human faculty.* London: Macmillan.

Gattegno, C. (1962). *Words in colour: Background and principles.* Reading, England: Educational Explorers.

Gattegno, C. (1969). *Reading with words in colour.* Reading, England: Educational Explorers.

Goodacre, E. (1971). *Children and learning to read.* London: Routledge & Kegan Paul.

Goodman, M., & Cundick, B. (1976). Learning rates with black and colored letters. *Journal of Learning Disabilities, 9,* 600-602.

Greene, J., & D'Oliveira, M. (1985). *Learning to use statistical tests in psychology.* Milton Keynes: Open University.

Hagen, J. (1967). The effect of distraction on selective attention. *Child Development, 38,* 685-694.

Hale, C., & Piper, R. (1973). Developmental trends in children's incidental learning: Some critical stimulus differences. *Developmental Psychology, 8,* 327-335.

Harrison, C. (1980). *Readability in the classroom.* Cambridge, England: Cambridge University.

Hinds, L. (1968). Studies in the use of colour. In Vilscek, E. (ed.) *A decade of innovations.* Newark, DE: International Reading Association.

Hochman, S. (1967). The effects of stress on Stroop color-word performance. *Psychonomic Science, 9,* 475-476.

Hochman, S. (1969). Stress and response competition in children's color-word performance. *Perceptual and Motor Skills, 28,* 115-118.

Howard, D. (1983). *Cognitive Psychology.* New York: Macmillan.

Huey, E. (1968). *The psychology and pedagogy of reading* (2nd ed.). New York: Macmillan.

Hunt, J., & Kirk, G. (1974). Criterion-referenced tests of school readiness: A paradigm with illustrations. *Genetic Psychology Monographs, 90,* 144-182.

Hunter, I. (1970). *Memory*. Baltimore, MD: Penguin.

Jensen, A., & Rohwer, W. Jr. (1966). The Stroop color-word test: A review. *Acta Psychologica, 25*, 36-93.

Jones, J. (1965). Colour as an aid to visual perception in early reading. *British Journal of Educational Psychology, 35*, 21-27.

Jones, J. (1967a). *Colour story reading*. London: Nelson.

Jones, J. (1967b). *Colour story reading: A research report*. London: Nelson.

Jones, J. (1968). Comparing i.t.a. with colour story reading. *Educational Research, 10*, 226-234.

Kanner, J. (1968). *The instructional effectiveness of color in television: A review of the evidence*. Washington, DC: DOA, Office of the Assistant Chief of Staff for Communications-Electronics. (ERIC Document Reproduction Service No. ED 015 675.)

Kintsch, W. (1982). *Memory and cognition*. Malabar, FL: Krieger.

Klare, G., Mabry, J., & Gustafson, L. (1955). The relationship of human interest to immediate retention and to acceptability of technical material. *Journal of Applied Psychology, 39*, 92-95.

Klatzky, R. (1980). *Human memory* (2nd ed.). New York: W.H. Freeman.

Klein, G. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology, 77*, 576-588.

Kroll, N., Kellicutt, M., Berrian, R., & Kreisler, A. (1974). Effects of irrelevant color changes on speed of visual recognition following short retention intervals. *Journal of Experimental Psychology, 103,* 97-106.

Lamberski, R., & Dwyer, F. (1981). The instructional effect of color in intermediate and delaved reaction. *Journal of Instructional Psychology, 8,* 122-131.

Lamberski, R., & Dwyer, F. (1983). The instructional effect of coding (color and black and white) on information acquisition and retrieval. *Educational Communication and Technology Journal, 31,* 9-21.

Lee, T. (1967). Writing the talking: An appraisal of words in colour. In Brown, A. *Reading: Current research and practice.* Edinburgh, Scotland: Chambers.

Light, L., & Berger, D. (1976). Are there long-term 'literal copies' of visually presented words? *Journal of Experimental Psychology: Human Learning and Memory, 2,* 654-662.

Lockmiller, P., & Di Nello, M. (1970). Words in color versus a basal reader program with retarded readers in grade 2. *Journal of Educational Research, 63,* 7.

Lyczak, R. (1976). Learning to read: The redundant cues approach. *Journal of Educational Psychology, 68,* 157-166.

McHugh, W. (1968). Reactions to using color in early reading instruction. In Vilscek, E. (ed.). *A decade of innovations.* Newark, DE: International Reading Association.

McNally, J., & Murray, W. (1962). *Key words to literacy.* London: Schoolmaster Publishing.

Moxon, C. (1962). *A remedial reading method.* London: Methuen.

Moyle, D., & Moyle, L. (1974). *Modern innovations in the teaching of reading.* London: University of London.

Norman, D. (1976). *Memory and attention* (2nd ed.). New York: Wiley.

Otto, W., & Askov, E. (1968). The role of color in learning and instruction. *Journal of Special Education, 2,* 155-165.

Park, D., & James, C. (1983). Effect of encoding instructions on children's spatial and color memory: Is there evidence for automaticity? *Child Development, 54,* 61-68.

Park, D., & Puglisi, J. (1985). Older adults' memory for the color of pictures and words. *Journal of Gerontology, 40,* 198-204.

Purcell, J. (1984). Using color to cue critical elements in word formation. *Hispania, 67,* 108-111.

Sabo, R., & Hagen, J. (1973). Color cues and rehearsal in short-term memory. *Child Development, 44,* 77-82.

Scarpino, T. (1972). The comparative effect of two strategies -- color versus black and white -- on achievement in college physical science (Doctoral dissertation, Pennsylvania State University). *Dissertation Abstracts International, 32,* 6222A.

Schiller, P. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology, 72,* 105-108.

Scott, J. (1968). Color as an aid in early reading instruction. In Vilscek, E. (ed.). *A decade of innovations.* Newark, DE: International Reading Association.

Sisson, C. (1968). Foreign language experience and color word interference. Report included in *Studies in Language and Language Behavior, Progress Report No. VI.* Ann Arbor, MI: University of Michigan (Center for Research on Language and Language Behavior).

Smith, J., & Watkins, H. (1972). *An investigation into some aspects of the illustration of primary school books.* Reading, England: University of Reading Typography Unit.

Southgate, V., & Roberts, G. (1970). *Reading -- Which approach?* London: University of London.

Spencer, H. (1969). *The visible word: Problems of legibility.* New York: Hastings House.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Tinker, M. (1963). *Legibility of print.* Ames, IA: Iowa State University Press.

Turner, A. (1984). *Color-coded vowels and spelling with visual cues in beginning reading.* NJ: Kean College of New Jersey. (ERIC Document Reproduction Service No. ED 244 226.)

Vernon, M. (1971a). *The psychology of perception* (2nd ed.). Harmondsworth, England: Penguin.

Vernon, M. (1971b). *Reading and its difficulties.* Cambridge, England: Cambridge University.

Waddington, M. (1965). Colour blindness in young children. *Educational Reasearch, 7,* 236-240.

Watts, L., & Nisbet, J. (1974). *Legibility in children's books.* Windsor, England: NFER.

Winters, J. Jr., & Hoats, D. (1984). Effects of isolation by color on mentally retarded and nonretarded persons' recall of printed words. *American Journal of Mental Deficiency, 89*, 310-312.

Table 1

Results of T Tests of Differences Between Related Colour (RC), Unrelated Colour (UC), and Black and White (BW) Test Scores for Memory and Spelling

Test	N	Mean	SD	T value	1-tail prob.
Memory					
RC	23	10.4783	3.566	1.85	0.036*
UC	23	8.5652	3.462		
UC	23	8.5652	3.462	-0.63	0.266
BW	24	9.2083	3.563		
RC	23	10.4783	3.566	1.22	0.114
BW	24	9.2083	3.563		
Spelling					
RC	23	7.4348	3.964	0.72	0.238
UC	23	6.6522	3.393		
UC	23	6.6522	3.393	-0.99	0.164
BW	24	7.6667	3.632		
RC	23	7.4348	3.964	-0.21	0.418
BW	24	7.6667	3.632		

* significant at .05 level

Table 2

Results of T Tests of Sex Differences for Pretest (P), Related Colour (RC), Unrelated Colour (UC), and Black and White (BW) Test Scores for Memory

Memory					
Test	Sex	N	Mean	SD	T value	1-tail prob.
P	M	35	8.7143	3.270	-1.99	.025*
	F	35	10.0286	3.899		
RC	M	11	10.3636	3.233	-0.15	.443
	F	12	10.5833	3.988		
UC	M	9	7.2222	2.333	-1.71	.051
	F	14	9.4286	3.857		
BW	M	15	8.2000	3.726	-2.05	.027*
	F	9	10.8889	2.667		

* significant at .05 level

Table 3

Results of T Tests of Sex Differences for Pretest (P), Related Colour (RC), Unrelated Colour (UC), and Black and White (BW) Test Scores for Spelling

		Spelling				
Test	Sex	N	Mean	SD	T value	1-tail prob.
P	M	33	6.4857	3.493	-2.17	.017*
	F	35	8.5429	4.375		
RC	M	11	7.7273	2.796	0.34	.369
	F	12	7.1667	4.914		
UC	M	9	4.8889	1.616	-2.53	.011* **
	F	14	7.7857	3.786		
BW	M	15	6.8667	3.461	-1.40	.091
	F	9	9.0000	3.708		

* significant at .05 level

** F value = 5.49 (2-tail prob. = 0.021)

Table 4

Results of the Spearman Rank Correlation Test Showing Correlations Between Chronological Age (CA) and Pretest (P), Related Colour (RC), Unrelated Colour (UC), and Black and White (BW) Test Scores for Memory and Spelling

Test	N	Spearman Coefficient	Significance
Memory			
P	70	.1115	.358
RC	23	.1567	.475
UC	23	.1449	.509
BW	24	.1504	.483
Spelling			
P	70	.1453	.230
RC	23	.2007	.358
UC	23	.0847	.701
BW	24	.2331	.273

Table 5

Results of the Spearman Rank Correlation Test Showing Correlations Between Reading Age (RA) and Pretest (P), Related Colour (RC), Unrelated Colour (UC), and Black and White (BW) Test Scores for Memory and Spelling

Test	N	Spearman Coefficient	Significance
Memory			
P	70	.5777	.000*
RC	23	.6813	.000*
UC	23	.6310	.001*
BW	24	.6614	.000*
Spelling			
P	70	.6274	.000*
RC	23	.6530	.001*
UC	23	.8002	.000*
BW	24	.6511	.001*

* significant at .001 level

Figure Captions

Figure 1. Pretest of neutral words printed in black and white.

Figure 2. The three stimulus conditions: 1) Colour words printed in related colours, 2) Colour words printed in unrelated colours, and 3) Colour words printed in black and white.

(Pretest)

kitten

jumper

dish

flour

blackberry

ribbon

umbrella

skirt

saucer

cheese

gloves

butter

dog

snow

spoon

corn

horse

forest

milk

rabbit

(Stimulus Condition 1)	(Stimulus Condition 2)	(Stimulus Condition 3)
purple (purple)*	purple (brown)	purple (b/w)
penny (brown)	penny (green)	penny (b/w)
grass (green)	grass (purple)	grass (b/w)
water (blue)	water (red)	water (b/w)
red (red)	red (green)	red (b/w)
beanstalk (green)	beanstalk (brown)	beanstalk (b/w)
dirt (brown)	dirt (blue)	dirt (b/w)
sea (blue)	sea (orange)	sea (b/w)
orange (orange)	orange (green)	orange (b/w)
tomato (red)	tomato (blue)	tomato (b/w)
bluebell (blue)	bluebell (red)	bluebell (b/w)
brown (brown)	brown (green)	brown (b/w)
ivy (green)	ivy (orange)	ivy (b/w)
ocean (blue)	ocean (red)	ocean (b/w)
strawberry (red)	strawberry (brown)	strawberry (b/w)
tea (brown)	tea (blue)	tea (b/w)
carrot (orange)	carrot (purple)	carrot (b/w)
violet (purple)	violet (brown)	violet (b/w)
moss (green)	moss (red)	moss (b/w)
cherry (red)	cherry (blue)	cherry (b/w)

* colours in parentheses indicate the colour of ink used.